Acta Psychologica 161 (2015) 7–17

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Retrieval monitoring is influenced by information value: The interplay between importance and confidence on false memory☆

Ian M. McDonough [a,*], Dung C. Bui [b], Michael C. Friedman [c], Alan D. Castel [d]

[a] Department of Psychology, University of Alabama, 505 Hackberry Lane, Tuscaloosa, AL 35487, USA
[b] Department of Psychology, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA
[c] Harvard Initiative for Learning and Teaching, Harvard University, 125 Mt. Auburn Street, Cambridge, MA 02138, USA
[d] Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 20 December 2014
Received in revised form 28 May 2015
Accepted 29 July 2015
Available online xxxx

Keywords:
False memory
Heuristics
Memory
Metamemory
Word recognition
Value

ABSTRACT

The perceived value of information can influence one's motivation to successfully remember that information. This study investigated how information value can affect memory search and evaluation processes (i.e., retrieval monitoring). In Experiment 1, participants studied unrelated words associated with low, medium, or high values. Subsequent memory tests required participants to selectively monitor retrieval for different values. False memory effects were smaller when searching memory for high-value than low-value words, suggesting that people more effectively monitored more important information. In Experiment 2, participants studied semantically-related words, and the need for retrieval monitoring was reduced at test by using inclusion instructions (i.e., endorsement of any word related to the studied words) compared with standard instructions. Inclusion instructions led to increases in false recognition for low-value, but not for high-value words, suggesting that under standard-instruction conditions retrieval monitoring was less likely to occur for important information. Experiment 3 showed that words retrieved with lower confidence were associated with more effective retrieval monitoring, suggesting that the quality of the retrieved memory influenced the degree and effectiveness of monitoring processes. Ironically, unless encouraged to do so, people were less likely to carefully monitor important information, even though people want to remember important memories most accurately.

1. Introduction

Every day people encounter more information than they can possibly remember. As technology advances, even more information is available through a variety of media outlets. This increased exposure to information heightens the need for people to selectively attend to and prioritize information based on their personal goals and motivations that are most important so that they will be more likely to later remember that information. Indeed, people are quite adept at later remembering information deemed to be important (e.g., Ariel, Dunlosky, & Bailey, 2009; Castel, Benjamin, Craik, & Watkins, 2002; Loftus & Wickens, 1970). Furthermore, people also expect to remember more important information than less important information (e.g., Festini, Hartley, Tauber, & Rhodes, 2013; Friedman & Castel, 2011; Kassam, Gilbert, Swencionis, & Wilson, 2009). While information deemed important might affect how people strategically attend to or encode information, whether the importance of information affects the strategies by which people search and evaluate their memories at retrieval (i.e., retrieval monitoring) is virtually unknown. Does event importance affect retrieval-monitoring processes? If so, what stages of retrieval monitoring are most affected? To address this question, the present study manipulated the perceived importance of information and the degree of retrieval monitoring.

Theories pertaining to retrieval monitoring propose that memorial expectations serve a key role in how people search memory and evaluate retrieved information (e.g., Gallo, 2010; Johnson, Hashtroudi, & Lindsay, 1993). For instance, people expect that visually distinctive events (i.e., events that contain many item-specific details) such as pictures will be better remembered than visually impoverished information such as words (e.g., Schacter & Wiseman, 2006). However, when the expected item-specific details cannot be retrieved for a particular event, people quickly and accurately reject the event as having occurred (e.g., "This item probably wasn't presented as a picture, because I'd remember more specific details."). In other words, people base their memory decisions on the expected qualitative characteristics of the to-be-recollected information and can avoid false memories if the retrieved information does not match those expectations. This type of retrieval-monitoring process has been referred to as a distinctiveness heuristic (Schacter, Israel, & Racine, 1999) and has been extended to conceptually distinctive relative to conceptually non-distinctive events as well (e.g., Gallo, Meadow, Johnson, & Foster, 2008; McDonough & Gallo, 2008). Thus, this reduced susceptibility to false memories following the

☆ Aspects of these data were presented at the 53rd Annual Meeting of the Psychonomic Society (2013, Toronto, Canada).

* Corresponding author at: 505 Hackberry Lane, Tuscaloosa, AL 35487, USA.
E-mail address: ian.mcdonough@gmail.com (I.M. McDonough).

http://dx.doi.org/10.1016/j.actpsy.2015.07.017
0001-6918/

encoding of distinctive events leads to increased monitoring effectiveness or the degree that retrieval monitoring leads to more accurate memory decisions.

Does important information leave more distinctive memory traces (i.e., contain more item-specific details) than less important information? If so, encoding important information might lead to fewer false memories when retrieved at a later point. Investigating how the importance of information affects memory has been approached using the value-directed-remembering (VDR) paradigm (Castel, 2007; Castel et al., 2002) in which participants study a list of words paired with a number or "value" (e.g., skate 7, cheek 12, fence 3) to remember for a later memory test. Prior to viewing the lists, participants are told that they will be awarded the point value associated with each remembered word, thus making words with higher numbers more valuable to remember than words with lower numbers (in the example above, remembering "cheek" would be more valuable than remembering "skate" or "fence"). VDR studies have demonstrated that participants recall or recognize more words that were paired with higher values than words with lower values (e.g., Ariel et al., 2009; Castel, Farb, & Craik, 2007; Castel et al., 2002; Loftus & Wickens, 1970). In addition, recent unpublished research indicates that high-value items are associated with more "remember" responses than low-value items, consistent with this idea (Cohen, Rissman, Harbert, Castel, & Knowlton, 2013; Hennessee, Cohen, Castel, & Knowlton, 2014).

While better memory and more "remember" responses for high-value items relative to low-value items is consistent with a distinctiveness account, these findings are not sufficient to argue that high-value items will lead to more effective retrieval monitoring than low-value items. For example, McDonough and Gallo (2008) showed that associating a presented object with a personal autobiographical memory resulted in both greater recognition memory and more frequent "remember" responses than judging whether an object was made in a factory, ostensibly due to the greater degree of elaboration during encoding. Consistent with retrieval-monitoring theories, searching memory for the autobiographical items was associated with fewer false memories than searching memory for factory items. Critically, repeating the factory judgments multiple times equated recognition memory and "remember" responses, but false memories continued to be lower when participants searched for the more distinctive types of events (i.e., autobiographical items). Using a different approach, Scimeca, McDonough, and Gallo (2011) showed that repeatedly presenting words led to greater recognition memory and more frequent "remember" responses than not repeating words, but false memories were *not* reduced when searching memory for repeated words compared with non-repeated words. They further showed that subjective ratings of item distinctiveness did not differ between repeated and non-repeated words. Together, these studies suggest that relative differences in memory strength and "remember" responses do not predict when retrieval monitoring will be effective, as measured through a reduction in false memories.

An alternative possibility is that important events lead people to try harder to remember the events by engaging in retrieval monitoring to a greater extent than less important events. That is, people might initiate multiple search attempts and take time to carefully evaluate the retrieved details, but nevertheless continue to be susceptible to false memories. This idea of monitoring engagement is orthogonal to the idea of monitoring effectiveness. For instance, people often take more time to respond when searching memory for non-distinctive events (i.e., more engagement), but the outcome of this additional effort often is in vain because memory accuracy is still poor (i.e., poor effectiveness; e.g., Gallo, Kensinger, & Schacter, 2006; Gallo, McDonough, & Scimeca, 2010).

Some experimental evidence is consistent with the idea that retrieval monitoring may not be affected by the importance of information. A study by Kassam et al. (2009) showed that people expected that knowing the importance of information both before and after encoding should impact subsequent memory, but they found that only knowing information before encoding impacted subsequent memory. This finding suggests that important information has large effects on attention and encoding strategies, but little or no effect on retrieval strategies. In addition, a recent study showed that processing information associated with more important information had the unintended consequence of creating more false memories relative to processing less important information when participants studied lists of semantic associates, known to create high levels of false memories (Bui, Friedman, McDonough, & Castel, 2013). While neither study isolated potential contributions of retrieval monitoring as a function of importance, these initial sources of evidence question the degree to which retrieval-monitoring processes effectively reduced susceptibility to false memories following the encoding of important information. In three experiments, we use well-established methods known to engage and manipulate retrieval-monitoring processes to reveal how those processes are affected by encoding information associated with different values.

1.1. Experiment 1

The primary aim of Experiment 1 was to test the degree to which encoding varying levels of importance affects retrieval monitoring. We used the criterial recollection task (Gallo, 2013; Gallo, Weiss, & Schacter, 2004)—a variant of traditional source memory tests—in which participants study different categories of stimuli at encoding (e.g., pictures and words), and then at test, they are oriented toward a specific category of stimuli (e.g., "Was this cue presented previously as a picture?"). To correctly respond, participants must recollect item-specific details for the queried format similar to source memory tests. In this way, criterial-recollection tests encourage retrieval monitoring of item-specific details across all items (e.g., Lindsay & Johnson, 1989; Multhaup, De Leonardis, & Johnson, 1999). However, unlike traditional source memory tests, criterial-recollection tests allow for the assessment of how effective these processes are during retrieval, as measured by the relative level of source misattributions across the different tests. That is, by having people assess each source separately, people can adjust their memorial expectations and subsequent retrieval-monitoring processes. This task has revealed different degrees of retrieval-monitoring effectiveness between many different types of stimuli (for review, see Gallo, 2013). Notably, this task uses unrelated words to minimize the effect of associative activation or the formation of gist traces (Brainerd & Reyna, 1998; Deese, 1959; Roediger & McDermott, 1995; Schacter, Verfaellie, & Pradere, 1996), and encourages retrieval monitoring (on average) for each category or stimulus type.

In Experiment 1, participants first studied a block of words (not paired with a value) to serve as lures during the criterial recollection tests (for a similar method, see Scimeca, McDonough, & Gallo, 2011).[1] Following this familiarization phase, participants studied words paired with low, medium, and high values (the value-study phase). Because targets (paired with values) and lures (not paired with values) were each only presented once, they should be relatively matched on familiarity. Participants then received three criterial-recollection tests: a low-value test, a medium-value test, and a high-value test. On each test, participants were oriented to one of the values and were asked whether the presented word was presented with the criterial value

[1] A modified version of the original criterial recollection task was used in Experiment 1. The original criterial recollection task has participants discriminate between the same two sources, but the class of items being searched for is different on each test. For example, cues for previously studied words and pictures would be presented on each test, but participants would either be asked if the cue was previously seen as a word on one test or as a picture on the other test. Thus, words studied in non-target values would serve as lures (e.g., words on the picture test). However, this approach leads to differential memory strength for lures across the different tests (i.e., picture lures on a word test would be more likely to be retrieved than word lures on a picture test). Instead, we adopted an approach used and validated by Scimeca et al. (2011) that implemented a familiarization phase first, which allowed lures to be identical across the three value tests. This modified version is conceptually similar to the original criterial recollection tests, but is more tightly controlled.